PPDG Achievements in the first 2 years.
Addendum to Renewal Proposal
PPDG-40

This is a list of accomplishments used as background material for the renewal proposal.

- Condor/Condor-G (early use in CMS, D0), SRM (STAR/JLAB), SRB (BaBar) and JLAB in general commitment to Web Service: all sustained efforts are noticeable since day 1 and all tools are in production and continuous strengthening and support (the projects are still alive, standardization and interoperability at all levels)
- Other tools showing constancy since year1 SciDAC: Clarens (CMS), SAMGrid and JIM (D0/CDF/…)
- There are mention of 3 special DOE news letter sent (in addition of the quarterly reports). From oldest to earliest
  . 2002: MOP (MOnte carlo distributed Production) in CMS (demo at SC2001), Atlas Magda, X509 PKI DOESG-CA, GDMP and Grid test beds.
  . News briefing was sent to DOE for the 50k CMS event generation on 5 sites Grid in June 2002
  . STAR use of HRM and GridFTP July 2002
- Formation at management level of the Trillium concept (collaboration cross project and outreach)
- Contributions highlights: Atlas work on monitoring passed to iGOC (used afterward in Grid2003)
- PPDG GT2.4 starts to include call back to VOMS
- DOE-SG and interoperability with EDG
- Lots of work with EDG, HEPCAL, WP-X, DataTAG seems like a worth mentioning PPDG deliverables compliance with Europe standards and sometimes, change in directions.
- ContributionS to the development and evolution of the GLUE schemaS (High Energy Physics Intergrid joint technical board). Amazing amount of development and strengthening as per the Globus products i.e. tools in production (from D0 use of DOE-SG and better support for CA, to STAR gridFTP problem resolution etc …) lead to consolidation of the basic bread-and-butter of the Grid: the GT. Shouldn't this be an achievement?? (or is it implicit ??)
- Grid test-beds and Atlas/CMS DC (leading to Grid2003 as per the evolution of the reports; should be emphasized)
- Monitoring seems to solidify around a pole of not only MDS but also a combination of Ganglia, MonaLisa and MDS. Adopted by many, does this come from the Trillium Monitoring project or the willingness by individual experiment to adopt the same scheme across facilities (the later is likely the answer) ??
- More recent work: DIAL, U-JDL, User Analysis framework on Grid (embryonic).
- Sociology issue: while early report mentioned a dire under-estimate of the task (lack of FTE, task complexity), experiments have brought extraneous manpower (externally to PPDG funded) to several projects. This should be viewed as an accomplishment (such an environment with manpower far below the magnitude

of the task would generally lead to a drop of the activities altogether in experiment. Maybe a s sign that one sees a greater benefit and worthwhile investment).

Doug's presentation: Security, Data management and Monitoring
- VOX: Globus gatekeeper was enhanced (standard in GT2.4) to permit an authentication, authorization callout to a local module to handle more elaborate decisions than is possible with the default behavior using a grid-mapfile
- GUMS: Grid User Management System, to integrate local site registration with VO registration
- DOE-Grid PKI Federation (ivDGL, GriPhyn, CEA, CERN, NIKHEF, SciDac, LBL, NERSC, Argonne, Science Grid, PPDG, Oakridge, BNL, LIP, SARA, CESNET, ESG, PPARC, INFN,). We deployed a certificate authority that enables scientific collaborations between US DOE, NSF and European Colleagues.
- SACRED, MyProxy, KCA, Smart Cards, VSC a product of recognizing problems with end-user management of long term private keys.
- File transfer: Bug fixes, feature enhancements in GridFTP client & server, as well as additions to gridftp protocol.
- RLS used by US ATLAS with simulation production run using the Chimera/Pegasus virtual data system from the GriPhyN project.
- BaBar using SRB for distributed data management and has interface to BaBar specific metadata (NCollDb)
- SRM interface with HPSS, Enstore, JASMine, DCache, CASTOR and any disk cache.
SRM used in STAR for robust file replication.
- Data Filtering and transformation: SOCATS. It is an STL based object caching and transport system. SOCATS is Grid-based middleware for efficient transfer of large relational datasets within he GAE framework. It uses Clarens for remote procedure call layer (SQL Database -> ODBC -> OTL -> SOCATS Server)
- Event selection manager used in STAR. Based on STACS and GCA client, it is integrated within the STAR analysis framework to extract selected events from ROOT files. SRM layer underneath, GRID-enabled, it is capable to move files across the Grid on demand.
- Phase II SBIR with Texh-X and SLAC: APIs for interactive analysis using the Grid.
- Web-Services and OGSI: JLab SOAP Client providing and FTP-like interface. Supported functions include list, mkdir, delete, copy. Another example would be the CMS Clarens Web-services layers.
- Monitoring: Globus MDS, MonaLisa and Ganglia deployed on many sites. Base tools used in Grid2003.
- GLUE Schema: Participation in standardizing information services schema between US and the European collaborators.

Jul-Sep 03
-   STAR/JLab started the definition of U-JDL
-   STAR adopted VDT
-   Joint effort on Grid2003 from PPDG team US-Atlas, US-CMS
-   Job management, monitoring and data management meeting started or re-started
-   SRM improvements (JLab)
-   BaBar SRB phased integration with the existing JImport.
-   D0 Jim installation finished, sandboxing of user input/output resolved, evaluation of VO
-   Data management: Replica Registration Service joint activity started.
-   SRB enhancement including disaster recovery, performance improvement and collection control by the local data grid

- RLS released in GT3.0, bug fix etc …
- Job management: new and more robust Condor version including Condor-G, gridshell; STAR/JLab U-JDL and service based on job submission portal; JIM worked with Condor to integrate Match Making.

Production Grids:
- STAR, JLab, BaBar data file transfer production activities continues; D0/CDF SAMGrid uses SAM data handling + Job and Information Management; CMS production Grid moved to Integrated Grid Test-bed (a.k.a. Grid3); Atlas developed a set of tools within the Grid3 context now turned to the personnel at the iGOC in Indiana.
- Data analysis WG: dataset definition, increased communication with HEPCAL LHC Computing Grid project.
- Caltech GAE workshop and agreement on the architectural components for a Grid Analysis Environment. GAE infrastructure document reshaped accordingly. CAIGEE use of Clarens portal for GAE client authentication and resource access. Clarens server demonstrated to operate with VOMS-generated certificates and interface to POOL metadata catalog.
- DIAL
- MonaLisa inclusion in VDT for Grid2003, deployed in STAR.
- Globus team work in Grid2003 (monitoring group). Glue schema, MDS scalability.

Site AAA
- Authorization callout for the Globus Toolkit in use at FNAL.
- BNL GUMS server integration within VOMS. Compatible with VO management tool for the LCG.
- Delocalized agent at SLAC, FNAL and BNL for the PPDG RA.

Mar-Jun 2003
(Mostly from the main achievements)
- Atlas: Athena reconstruction using Chimera in collaboration with GriPhyn, iVDGL. Globus RLS tested within Chimera context.
- BaBar: BdbServer++ for TierA to TierC distribution to use SRB BaBar Metadata catalog. Use of SRB in production.
- CMS: Grid Analysis software/ authorization in Clarens d.a.f. . SOCAT for large relational datasets within GAE framework. MCRunjob integrated.
- Condor: JIM deployment. DAGMAN and dynamic DAG support.
- D0: deployment of JIM in production for D0/CDF. Sandboxing.
- Globus: GT2.4 and release of GT3 Beta in June. Modified Globus Tool security lib for callout to local VO Authorization service. Work done within PPDG-SiteAAA WG. Callout support for VOMS. Support CAS as a VO server at LBNL.
- JLab: SRM v2.x development and testing. Production level files replication between several sites.
- SRB: version 2.0.0 and 2.1.0. Support for // IO streams, bulk registration, bulk data access, metadata access control. WSDL/SOAP interface. Integration of GSI. Overall improvements (checksums, status, file transport). Work with BaBar.
- SRM: HRM between BNL and LBNL. SRM v2.x specification. GGF BOF.

- STAR: Production version of Job-Scheduler. Condor-G job scheduling. Monitoring systems.

Jan-Mar 2003
- Interaction with GGF, HEP Intergrid JTB continues
- Testing in joint projects for Middleware and Glue schema.
- SciDac presentation
- PPDG project and participant at the Globus World (9 presentations), CHEP03 (3 plenary, 26 parallel, 6 posters)
- DataTAG, LHC, EDG layering middleware on VDT which in turns is an extended benefit to all PPDG teams but also other application domains using core Globus, Condor and SRM.
- PPDG joined GAG (lhc Grid Application Group)
- Glue Schema work and CE used in EDG and part of MDS

Oct-Dec-2002
- SC2002 demo presentations: World Grid interoperability.
- Workshop on Analysis tools at Caltech
- Workshop on Troubleshooting and diagnostic in Grid Environments (35 participants)
- Site-AA project completed its scheduled work and reported to PPDG Steering and the funding MICS (Marie-Anne that is)
- SRM workshop at CERN in December: further define functionality and standardize the interface of SRM (with LBNL, FermiLab, JLab, Data Grid WP2 and WP5). New functional design document.
- SAM and JIM: prototype of JIM (D0) shown at SC2002 (Baltimore). Some level of brokering, web based monitoring of grid system and jobs. Use of ClassAd's, Condor-G, GRAM, GSI , etc … use of FermiLab Kerberos principles.
- Information Providers based on Glue CE used as part of the WorldGrid SC demos.
- IEPM (Performance Monitoring). Mention of Web-Service at this stage i.e. PingER results available via WS (instead of sole MDS/Globus). Work with EDG and NASA (integration of feedback). PingER use in ESnet, ICFA, SCIC, ICTP
- Intensified use of HRM in STAR environment between BNL/LBNL. FMT tested. WSDL-based HRM Version 1.1 compatible with Fermilab and JLab. Use of gSOAP instead of GLUE (used by Fermi and JLab). Lack of compatibility found. DRM-NeST integration done. Work at CERN on SRM v2.0.
- Globus RFT as first OGSA service.
- BaBar demonstration of data replication from SLAC to IN2P3 made at SC2002. File registration in bulk.
- Site-AAA: proxy generation, common authorization callout and gridFTPd. Incident handling.
- Globus CAS evaluation.
- Usage of Magda in Atlas in production for DC1. BNL HPSS to CERN Castor.
- Job Scheduling in STAR deployed at both sites, Condor-G started.
- GDMP testing in CMS

- JLab interoperable SRM demo at SC2002 (with Fermi Enstore). Transferred data to Glasgow with Jasmine-base SRM.
- Grid interface with interactive Analysis Tools. Grid enabled of the distributed JAS analysis system. Remote analysis in CMS via Clarens servers (US sites).

Jul-sep 2002
- One news briefing to report achievements of STAR use of HRM and GridFTP for production data replication and storage ; significant advances in Atlas, CMS, D0, BaBar and JLAB test beds and grid use.
- Issue pointed out as per the deployment and later support of grid middleware now in production use in experiment.
- Common JDL work by Andy
- Joint PPDG, GriPhyn and iVDGL Monitoring Project
- DOE CA in action (479 certificates)
- Atlas: Grappa, GDMP, Magda
- STAR: mention of the Scheduler, Monitoring work
- Other section are pretty much saying the same thing than quarterly +1

Apr-June 2002
End of the first year of the SciDac funded PPDG CP. Focus on peer experiment-computer science group. Projected sharing of technology, interface to common services and move up the service stack.
- News briefing was sent to DOE for the 50k CMS event generation on 5 sites Grid (less that 10% of CMS simulation, was accepted as an official CMS simulation)
- Troubleshooting identified as important. Submitted pre-proposal to DOE for resolving the topic of error and recovery.
- Common services, application Grid service components refocus
- Increased collaboration with HEP+NP projects, Trillium iVDGL, GripHyn and PPDG. The cross collaboration has lead to a concept of Trillium. Allows for common and convergent goals but also movement of resources between projects as needed.
- PPDG contribution to VDT.
- Collaboration with EDG and EDT (European DataTAG). DataTAG through GLUE iVDGL/DataTAG.
- High Energy Physics Intergrid joint technical board sponsoring the GLUE schema document(ation) (Doug, Ruth, Peter).
- Work with EDG for GDMP subscribe/subscribe and support for a Master Replica.
- Hardening of Condor and Globus
- DAGMAN
- Pacman cache use in Atlas for DC1
- SRM file transfer monitoring tool (FMT) in STAR (BNL/PDSF) developed by SDM
- Globus 2.0 and hanging transfer resolved in that release (STAR-DDM LBNL/SDM) and use of bbftp.
- Planned Gobus RFT OGSA service
- BaBar SRB and interoperation with AFS tokens. Switching to GSI.

- JASMine and SRB
- Integration of gridFTP started at 3 sites (FNAL, JLAB and SLAC)
- Use of third party VO management tool
- DIAL and distributed analysis
- Clarens robustness and improvements in CMS

Jan-Mar 2002
- First inter-continental file transfer using DOE-SG demonstrated in D0 using the SAM system
- BaBar European sites adopted SRB for dataset replication
- Physics analysis integrated with Grid services essential ingredient of requirements
- BaBar SRB and MCAT
- DOESG since January 2002 (RA and CA).
- EDG test bed interface using DOE-SG certificate
- Clarens in CMS
- D0 work with Condor-G/GRAM and Globus version incompatibility lead to a better support for alternate CA.
- Web Service in JLAB for Replica Catalog, SRM and File Transfer Service
- DRM/SRM by SDM strengthening
- SRB WSDL interface demo at GGF

Oct-dec 2001
- News update on Januay 29th 2002: MOP (MOnte carlo distributed Production) in CMS (demo at SC2001), Atlas Magda, X509 PKI DOESG-CA, GDMP.
- JDL at different stage of sophistication : CONDOR, Globus GRAM, SAM, CMS
- Collaboration between SRM SciDAC project and PPDG: delivery of a standard interface specification.
- Work with other projects (GriPhyn, iVDGL) and beginning of discussion on joint effort.
- D0-Condor, CMS-Condor
- BaBar SRB
- More of GDMP, RLS etc …

Jul-sep 2001
- GDMP, SRB, Condor are the main activities (in development)
- JLAB XML based interface and Web-Service direction
- BaBar SRB day1 investment
- SDM center SRM/TRM/HRM